

Search History

SCORE Search Results Details for Application 09734329 and Search Result us-09-734-329b-2.p2n.rni.

Score Home	Retrieve Application	SCORE System	SCORE	Comments /
Page	List	Overview	FAQ	Suggestions

This page gives you Search Results detail for the Application 09734329 and Search Result us-09-734-329b-2.p2n.rni.

[start](#)

[Go Back to previous](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2006, 09:19:25 ; Search time 269 Seconds
(without alignments)
4465.620 Million cell updates/sec

Title: US-09-734-329B-2
Perfect score: 2384
Sequence: 1 MASSLLEEEAHYGSSPLAML.....PAPPEKAHGGSPSEQSNLLEI 428

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US09734329/runat_25072006_091908_3760/app_query.fasta_1
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abss03h -USER=US09734329_CGN_1_1_204_@runat_25072006_091908_3760
-NCPU=6 -ICPU=3 -NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : Issued_Patents_NA:*
1: /EMC_Celerra_SIDS3/ptodata/2/ina/1_COMB.seq:*

2: /EMC_Celerra_SIDS3/ptodata/2/ina/5_COMB.seq:*
 3: /EMC_Celerra_SIDS3/ptodata/2/ina/6A_COMB.seq:*
 4: /EMC_Celerra_SIDS3/ptodata/2/ina/6B_COMB.seq:*
 5: /EMC_Celerra_SIDS3/ptodata/2/ina/7_COMB.seq:*
 6: /EMC_Celerra_SIDS3/ptodata/2/ina/H_COMB.seq:*
 7: /EMC_Celerra_SIDS3/ptodata/2/ina/PCTUS_COMB.seq:*
 8: /EMC_Celerra_SIDS3/ptodata/2/ina/PP_COMB.seq:*
 9: /EMC_Celerra_SIDS3/ptodata/2/ina/RE_COMB.seq:*
 10: /EMC_Celerra_SIDS3/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	818.5	34.3	3471	4	US-10-094-749-1170	Sequence 1170, Ap
2	505	21.2	3289	3	US-09-620-312D-436	Sequence 436, App
3	459.5	19.3	2984	3	US-09-949-016-4299	Sequence 4299, Ap
4	453	19.0	3937	3	US-10-164-595-7	Sequence 7, Appli
5	449.5	18.9	3722	3	US-10-164-595-9	Sequence 9, Appli
6	449.5	18.9	3862	3	US-10-164-595-5	Sequence 5, Appli
7	449.5	18.9	3985	3	US-10-164-595-3	Sequence 3, Appli
8	437.5	18.4	1430	3	US-09-492-985-1	Sequence 1, Appli
9	437.5	18.4	1471	3	US-09-492-985-11	Sequence 11, Appl
10	424.5	17.8	1057	3	US-09-949-016-4554	Sequence 4554, Ap
11	418.5	17.6	1079	3	US-10-131-827-8867	Sequence 8867, Ap
12	402	16.9	4863	3	US-09-949-016-3725	Sequence 3725, Ap
13	393	16.5	4859	3	US-09-814-915A-34	Sequence 34, Appl
14	393	16.5	4859	3	US-09-949-016-168	Sequence 168, App
15	393	16.5	4859	4	US-09-880-107-1647	Sequence 1647, Ap
16	366.5	15.4	589	3	US-09-229-037-29	Sequence 29, Appl
17	366.5	15.4	589	3	US-09-478-681-29	Sequence 29, Appl
18	366.5	15.4	589	3	US-09-779-233-17	Sequence 17, Appl
19	366.5	15.4	589	3	US-09-706-243A-29	Sequence 29, Appl
20	366.5	15.4	589	3	US-09-942-087A-29	Sequence 29, Appl
21	366.5	15.4	589	4	US-09-897-844A-29	Sequence 29, Appl
22	366.5	15.4	589	4	US-10-412-105-17	Sequence 17, Appl
23	366.5	15.4	589	5	US-09-844-508-28	Sequence 28, Appl
24	366.5	15.4	589	5	US-10-245-415B-29	Sequence 29, Appl
25	360.5	15.1	1734	3	US-09-949-016-2024	Sequence 2024, Ap
26	359	15.1	298	3	US-09-229-037-14	Sequence 14, Appl
27	359	15.1	298	3	US-09-478-681-14	Sequence 14, Appl
28	359	15.1	298	3	US-09-779-233-2	Sequence 2, Appli
29	359	15.1	298	3	US-09-706-243A-14	Sequence 14, Appl
30	359	15.1	298	3	US-09-942-087A-14	Sequence 14, Appl
31	359	15.1	298	4	US-09-897-844A-14	Sequence 14, Appl
32	359	15.1	298	4	US-10-412-105-2	Sequence 2, Appli
33	359	15.1	298	5	US-09-844-508-10	Sequence 10, Appl
34	359	15.1	298	5	US-10-245-415B-14	Sequence 14, Appl
35	353.5	14.8	1889	2	US-08-946-241B-1	Sequence 1, Appli
36	353.5	14.8	1889	2	US-08-946-241B-8	Sequence 8, Appli
37	353.5	14.8	1889	3	US-09-309-053-1	Sequence 1, Appli
38	353.5	14.8	1889	3	US-09-309-053-8	Sequence 8, Appli
39	353.5	14.8	2639	3	US-10-087-641-2	Sequence 2, Appli
40	353.5	14.8	2639	4	US-10-087-346-2	Sequence 2, Appli
41	351.5	14.7	1953	3	US-09-814-915A-85	Sequence 85, Appl
42	351	14.7	292	3	US-09-395-448-14	Sequence 14, Appl

43	351	14.7	292	3	US-09-925-796-14	Sequence 14, Appl
44	351	14.7	292	3	US-09-941-450-14	Sequence 14, Appl
45	351	14.7	298	3	US-09-229-037-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-10-094-749-1170
; Sequence 1170, Application US/10094749
; Patent No. 6979557
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOUYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1170
; LENGTH: 3471
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-1170

Alignment Scores:

Pred. No.:	5.34e-39	Length:	3471
Score:	818.50	Matches:	197
Percent Similarity:	53.6%	Conservative:	51
Best Local Similarity:	42.5%	Mismatches:	152
Query Match:	34.3%	Indels:	63
DB:	4	Gaps:	18

US-09-734-329B-2 (1-428) x US-10-094-749-1170 (1-3471)

Qy	1	MetAlaSerSerLeuLeuGluGluGluAlaHisTyrGlySerSerProLeuAlaMetLeu	20
		::: :::	
Db	89	ATGGCAACTTCACTTCTAGGGGAAGAACCGAGGTTGGGATCGACTCCTCTGGCCATGCTT	148
Qy	21	ThrAlaAlaCysSerLysPheGlyGlySerSerPro-----LeuArgAsp---	35

SCORE Search Results Details for Application 09734329 and Search Result us-09-734-329b-2.p2n.rnpbm.

Score Home Page	Retrieve Application List	SCORE System Overview	SCORE FAQ	Comments / Suggestions
---------------------------------	---	---------------------------------------	---------------------------	--

This page gives you Search Results detail for the Application 09734329 and Search Result us-09-734-329b-2.p2n.rnpbm.

[start](#)

[Go Back to previous](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2006, 09:20:29 ; Search time 1639 Seconds
(without alignments)
4813.094 Million cell updates/sec

Title: US-09-734-329B-2
Perfect score: 2384
Sequence: 1 MASSLLEEEAHYGSSPLAML.....PAPPEKAHGGSPSEQSNLLEI 428

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US09734329/runat_25072006_091917_3923/app_query.fasta_1
-DB=Published_Applications_NA_Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss02h
-USER=US09734329 @CGN_1_1_1675 @runat_25072006_091917_3923 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA_Main:*

```

1: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US08_PUBCOMB.seq:*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
9: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
10: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
11: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
12: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
13: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
14: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11B_PUBCOMB.seq:*
15: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11C_PUBCOMB.seq:*
16: /EMC_Celerra_SIDS3/ptodata/2/pubpna/US11D_PUBCOMB.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2384	100.0	2960	3	US-09-734-329-1	Sequence 1, Appli
2	818.5	34.3	3471	7	US-10-094-749-1170	Sequence 1170, Ap
3	686	28.8	2076	13	US-11-097-143-848	Sequence 848, App
4	676	28.4	1675	10	US-10-450-763-15125	Sequence 15125, A
5	671.5	28.2	1206	7	US-10-029-386-20198	Sequence 20198, A
6	650.5	27.3	4311	13	US-11-097-143-847	Sequence 847, App
7	553.5	23.2	1830	7	US-10-312-680-1	Sequence 1, Appli
8	551.5	23.1	1197	7	US-10-312-680-3	Sequence 3, Appli
9	505	21.2	2862	8	US-10-479-435-50	Sequence 50, Appl
10	505	21.2	3024	10	US-10-888-875-1	Sequence 1, Appli
11	505	21.2	3289	6	US-10-037-270-436	Sequence 436, App
12	505	21.2	3289	7	US-10-117-722-436	Sequence 436, App
13	505	21.2	3289	10	US-10-122-851-436	Sequence 436, App
14	493.5	20.7	3187	13	US-11-097-143-11864	Sequence 11864, A
c 15	493.5	20.7	5187	13	US-11-097-143-11863	Sequence 11863, A
16	479.5	20.1	2475	16	US-11-136-527-2078	Sequence 2078, Ap
17	453	19.0	3936	10	US-10-888-875-3	Sequence 3, Appli
18	453	19.0	3937	10	US-10-717-665-7	Sequence 7, Appli
19	449.5	18.9	3709	6	US-10-044-090-360	Sequence 360, App
20	449.5	18.9	3722	10	US-10-717-665-9	Sequence 9, Appli
21	449.5	18.9	3862	10	US-10-717-665-5	Sequence 5, Appli
22	449.5	18.9	3985	10	US-10-717-665-3	Sequence 3, Appli
23	438	18.4	4285	6	US-10-012-600B-57	Sequence 57, Appl
24	422	17.7	600	16	US-11-136-527-6174	Sequence 6174, Ap
25	418.5	17.6	1079	7	US-10-131-827-8867	Sequence 8867, Ap
26	414.5	17.4	1935	13	US-11-097-143-38099	Sequence 38099, A
c 27	414	17.4	3388	13	US-11-097-143-13342	Sequence 13342, A
28	410.5	17.2	1133	10	US-10-956-157-2550	Sequence 2550, Ap
29	408	17.1	1615	7	US-10-264-049-704	Sequence 704, App
30	393.5	16.5	882	3	US-09-764-864-642	Sequence 642, App
31	393.5	16.5	900	3	US-09-764-864-211	Sequence 211, App
32	393	16.5	4859	3	US-09-954-456-493	Sequence 493, App
33	393	16.5	4859	3	US-09-880-107-1647	Sequence 1647, Ap
34	393	16.5	4859	6	US-10-171-311-21	Sequence 21, Appl
35	393	16.5	4859	7	US-10-172-118-572	Sequence 572, App

```

RESULT 1
US-09-734-329-1
; Sequence 1, Application US/09734329
; Patent No. US20020156031A1
; GENERAL INFORMATION:
; APPLICANT: de CROMBRUGGHE, BENOIT
; APPLICANT: NAKASHIMA, KAZUHISA
; APPLICANT: ZHOU, XIN
; TITLE OF INVENTION: MASTER BONE FORMATION TRANSCRIPTION FACTOR:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: UTXC:666
; CURRENT APPLICATION NUMBER: US/09/734,329
; CURRENT FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2960
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (100)..(1383)
US-09-734-329-1

```

Pred. No.:	2.31e-196	Length:	2960
Score:	2384.00	Matches:	428
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	3	Gaps:	0

Qy	1	MetAlaSerSerLeuLeuGluGluGluAlaHisTyrGlySerSerProLeuAlaMetLeu	20
Db	100	ATGGCGTCCTCTCTGCTTGAGGAAGAAGCTCACTATGGCTCCAGTCCCCTGGCCATGCTG	159
Qy	21	ThrAlaAlaCysSerLysPheGlyGlySerSerProLeuArgAspSerThrThrLeuGly	40
Db	160	ACTGCAGCCTGCAGCAAATTTGGCGGCTCTAGCCCTCTGCGGGACTCAACAACCCTGGGG	219
Qy	41	LysGlyGlyThrLysLysProTyrAlaAspLeuSerAlaProLysThrMetGlyAspAla	60
Db	220	AAAGGAGGCACAAAGAAGCCATACGCTGACCTTTCAGCCCCCAAACCATGGGGGACGCC	279
Qy	61	TyrProAlaProPheSerSerThrAsnGlyLeuLeuSerProAlaGlySerProProAla	80

SCORE Search Results Details for Application 09734329 and Search Result us-09-734-329b-2.p2n.rnpbn.

[Score Home Page](#)
[Retrieve Application List](#)
[SCORE System Overview](#)
[SCORE FAQ](#)
[Comments / Suggestions](#)

This page gives you Search Results detail for the Application 09734329 and Search Result us-09-734-329b-2.p2n.rnpbn.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2006, 09:24:46 ; Search time 430 Seconds
(without alignments)
2144.788 Million cell updates/sec

Title: US-09-734-329B-2
Perfect score: 2384
Sequence: 1 MASSLLEEEAHYGSSPLAML.....PAPPEKAHGGSPSEQSNLLEI 428

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 899801 seqs, 718270062 residues

Total number of hits satisfying chosen parameters: 1799602

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US09734329/runat_25072006_091919_3956/app_query.fasta_1
-DB=Published_Applications_NA_New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -HOST=abss03h
-USER=US09734329_@CGN_1_1_150_@runat_25072006_091919_3956 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : Published_Applications_NA_New:*

1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq:*
 2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq:*
 3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq:*
 5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
 6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq:*
 7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq:*
 8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq1:*
 9: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	674	28.3	3794	8	US-11-266-748A-31703	Sequence 31703, A
2	604.5	25.4	970	8	US-11-266-748A-92977	Sequence 92977, A
c 3	604.5	25.4	970	8	US-11-266-748A-145788	Sequence 145788, A
4	551.5	23.1	1812	8	US-11-266-748A-28121	Sequence 28121, A
5	505	21.2	2655	8	US-11-266-748A-28794	Sequence 28794, A
6	505	21.2	7667	7	US-11-317-330A-18	Sequence 18, Appl
c 7	459.5	19.3	1678	8	US-11-266-748A-262860	Sequence 262860, A
8	459.5	19.3	1678	8	US-11-266-748A-323377	Sequence 323377, A
9	453	19.0	3979	8	US-11-266-748A-24836	Sequence 24836, A
10	449.5	18.9	1135	8	US-11-266-748A-193673	Sequence 193673, A
11	449.5	18.9	1135	8	US-11-266-748A-227203	Sequence 227203, A
12	449.5	18.9	2217	8	US-11-266-748A-29033	Sequence 29033, A
13	449.5	18.9	2529	8	US-11-266-748A-27455	Sequence 27455, A
14	449.5	18.9	3920	8	US-11-266-748A-29348	Sequence 29348, A
15	447.5	18.8	1522	8	US-11-266-748A-353547	Sequence 353547, A
16	447.5	18.8	1522	8	US-11-266-748A-384508	Sequence 384508, A
c 17	447.5	18.8	1522	8	US-11-266-748A-436926	Sequence 436926, A
18	445	18.7	1000	8	US-11-266-748A-284152	Sequence 284152, A
c 19	445	18.7	1000	8	US-11-266-748A-335581	Sequence 335581, A
20	438	18.4	2888	8	US-11-145-307A-253	Sequence 253, App
21	438	18.4	3031	8	US-11-266-748A-56359	Sequence 56359, A
22	438	18.4	3288	8	US-11-266-748A-28709	Sequence 28709, A
c 23	401.5	16.8	2961	8	US-11-266-748A-80279	Sequence 80279, A
24	401.5	16.8	2961	8	US-11-266-748A-133090	Sequence 133090, A
25	399.5	16.8	1000	8	US-11-266-748A-291349	Sequence 291349, A
c 26	399.5	16.8	1000	8	US-11-266-748A-342778	Sequence 342778, A
27	393.5	16.5	842	8	US-11-266-748A-363880	Sequence 363880, A
28	393.5	16.5	842	8	US-11-266-748A-388190	Sequence 388190, A
c 29	393.5	16.5	842	8	US-11-266-748A-447259	Sequence 447259, A
30	393	16.5	4859	8	US-11-266-748A-28713	Sequence 28713, A
31	392	16.4	2129	8	US-11-293-697-1978	Sequence 1978, Ap
32	371	15.6	1000	8	US-11-266-748A-403925	Sequence 403925, A
c 33	371	15.6	1000	8	US-11-266-748A-474971	Sequence 474971, A
34	370.5	15.5	1618	8	US-11-266-748A-29737	Sequence 29737, A
35	369.5	15.5	499	8	US-11-266-748A-141	Sequence 141, App
36	364.5	15.3	1152	6	US-10-511-937-402	Sequence 402, App
37	358	15.0	1668	8	US-11-266-748A-358069	Sequence 358069, A
c 38	358	15.0	1668	8	US-11-266-748A-441448	Sequence 441448, A
39	354.5	14.9	1869	8	US-11-293-697-1940	Sequence 1940, Ap
40	353.5	14.8	2639	8	US-11-235-814-2	Sequence 2, Appli
41	351.5	14.7	1953	8	US-11-266-748A-32031	Sequence 32031, A
42	346.5	14.5	1953	8	US-11-266-748A-186388	Sequence 186388, A

	43	345	14.5	2908	8	US-11-235-814-1	Sequence 1, Appli
	44	342.5	14.4	1000	8	US-11-266-748A-117354	Sequence 117354,
c	45	342.5	14.4	1000	8	US-11-266-748A-159518	Sequence 159518,

ALIGNMENTS

RESULT 1

US-11-266-748A-31703
; Sequence 31703, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31703
; LENGTH: 3794
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-31703

Alignment Scores:

Pred. No.:	3.09e-20	Length:	3794
Score:	674.00	Matches:	176
Percent Similarity:	49.3%	Conservative:	43
Best Local Similarity:	39.6%	Mismatches:	131
Query Match:	28.3%	Indels:	96
DB:	8	Gaps:	20

US-09-734-329B-2 (1-428) x US-11-266-748A-31703 (1-3794)

Qy	18	AlaMetLeuThrAlaAlaCysSerLysPheGlyGlySerSerProLeuArgAspSerThr	37
Db	276	GCAATGCTAACCGCTGTCTGCGGC-----	299
Qy	38	ThrLeuGlyLysGlyGlyThrLysLysProTyrAlaAspLeuSerAlaProLysThrMet	57
		:::	
Db	300	TCTCTGGGCAGCCAGCACACGGAAGCGCCGACGCTCCCC-GCCGCGCCTCGA-----	352

Comments /
Suggestions

[Go Back to previous page](#)

<http://es/ScoreAccessWeb/GetItem.action?AppId=09734329&seqId=568837&ItemName=us-...> 8/1/06

2: geneseqn1990s:*
 3: geneseqn2000s:*
 4: geneseqn2001as:*
 5: geneseqn2001bs:*
 6: geneseqn2002as:*
 7: geneseqn2002bs:*
 8: geneseqn2003as:*
 9: geneseqn2003bs:*
 10: geneseqn2003cs:*
 11: geneseqn2003ds:*
 12: geneseqn2004as:*
 13: geneseqn2004bs:*
 14: geneseqn2005s:*
 15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

SUMMARIES							
Result	%		Query			Description	
	No.	Score	Match	Length	DB ID		
c	1	2384	100.0	2960	6	AAD40709	Aad40709 Mouse ost
	2	2264.5	95.0	3153	13	ADR08250	Adr08250 Full leng
	3	854.5	35.8	1737	12	ADP28826	Adp28826 Human sec
	4	836.5	35.1	3128	13	ADR07977	Adr07977 Full leng
	5	818.5	34.3	3471	10	ADA53602	Ada53602 Human cod
	6	818.5	34.3	3471	14	AEB53898	Aeb53898 DNA encod
	7	774.5	32.5	1473	12	ADO00467	Ado00467 Novel hum
	8	774.5	32.5	1473	12	ADN98898	Adn98898 Novel hum
	9	686	28.8	2076	4	ABL02405	Abl02405 Drosophil
	10	676	28.4	1675	5	AAS79321	Aas79321 DNA encod
11	674	28.3	3794	13	ADR07645	Adr07645 Full leng	
12	671.5	28.2	1206	12	ACH87003	Ach87003 Human gen	
13	669	28.1	3846	13	ADR07511	Adr07511 Full leng	
14	650.5	27.3	4311	4	ABL02404	Abl02404 Drosophil	
15	553.5	23.2	1830	6	ABA96948	Aba96948 Mouse Spl	
16	552	23.2	1925	14	ADX83136	Adx83136 Human TEG	
17	551.5	23.1	1197	6	ABA96949	Aba96949 Human Spl	
18	551.5	23.1	1554	10	ADC30435	Adc30435 Human nov	
19	505	21.2	2862	10	AAD51694	Aad51694 Human nuc	
20	505	21.2	3024	14	ADZ72114	Adz72114 Human tra	
21	505	21.2	3289	4	AAI58549	Aai58549 Human pol	
22	505	21.2	3289	5	ADQ98766	Adq98766 DNA encod	
23	505	21.2	3289	9	ADB48526	Adb48526 Novel hum	
24	501.5	21.0	3090	4	AAI60335	Aai60335 Human pol	
25	493.5	20.7	3187	4	ABL09749	Abl09749 Drosophil	
c	26	493.5	20.7	5187	4	ABL09748	Abl09748 Drosophil
27	453	19.0	3936	14	ADZ72116	Adz72116 Human tra	
28	453	19.0	3937	10	ADK65768	Adk65768 Angiogene	
29	449.5	18.9	1848	13	ADT07445	Adt07445 Human col	
30	449.5	18.9	1923	13	ADT07444	Adt07444 Human col	
31	449.5	18.9	3312	13	ADT07441	Adt07441 Human col	
32	449.5	18.9	3430	13	ADT07439	Adt07439 Human col	
33	449.5	18.9	3448	13	ADT07440	Adt07440 Human col	
34	449.5	18.9	3548	6	ABK84634	Abk84634 Human cDN	
35	449.5	18.9	3548	14	ADX07998	Adx07998 Cyclin-de	
36	449.5	18.9	3548	15	AEF74819	Aef74819 Human pol	
37	449.5	18.9	3707	13	ADT07438	Adt07438 Human col	

38	449.5	18.9	3709	8	ABX63360	Abx63360 Human cDN
39	449.5	18.9	3722	10	ADK65770	Adk65770 Angiogene
40	449.5	18.9	3763	14	AEA19695	Aea19695 Novel hum
41	449.5	18.9	3862	10	ADK65766	Adk65766 Angiogene
42	449.5	18.9	3920	13	ADR83463	Adr83463 Human Sp3
43	449.5	18.9	3985	10	ADK65764	Adk65764 Angiogene
44	438	18.4	2888	15	AEe80965	Aee80965 Human cDN
45	438	18.4	3288	6	ABK83634	Abk83634 Human cDN

ALIGNMENTS

RESULT 1

AAD40709

ID AAD40709 standard; cDNA; 2960 BP.

XX

AC AAD40709;

XX

DT 30-OCT-2002 (first entry)

XX

DE Mouse osterix cDNA.

XX

KW Bone formation; transcription factor; osteoblast; Paget's disease;
KW osterix protein; glucocorticoid; osteoporosis; cytostatic; cytokine;
KW periodontal disease; tooth loss; bone fracture; rheumatoid arthritis;
KW metastatic bone disease; gene therapy; growth factor; osteopathic;
KW differentiation; mouse; chromosome 15; gene; ss.

XX

OS Mus musculus.

XX

FH Key Location/Qualifiers

FT CDS 100. .1386

FT /*tag= a

FT /product= "Mouse osterix protein"

XX

PN WO200244380-A2.

XX

PD 06-JUN-2002.

XX

PF 30-NOV-2001; 2001WO-US044898.

XX

PR 30-NOV-2000; 2000US-00734329.

XX

PA (TEXA) UNIV TEXAS SYSTEM.

XX

PI De Crombrugghe B, Nakashima K, Zhou X;

XX

DR WPI; 2002-519587/55.

DR P-PSDB; AAE25076.

XX

PT Novel DNA segment encoding Osterix polypeptide which is a master
PT transcription factor that controls osteoblast differentiation and is
PT useful for treating osteoporosis, in patient by stimulating bone
PT formation.

XX

PS Claim 17; Page 138-141; 144pp; English.

XX

CC The invention relates to a master bone formation transcription factor
CC that controls osteoblast differentiation, osterix protein and its
CC corresponding nucleic acid sequence. Osterix protein and its DNA and
CC agents that interact with the protein to activate or stimulate the

CC differentiation of bone cells is used for treatment of glucocorticoid
 CC induced osteoporosis, Paget's disease, periodontal disease, tooth loss,
 CC bone fractures, rheumatoid arthritis, metastatic bone disease, etc.
 CC Osterix DNA is useful as probes or primers in nucleic acid hybridisation
 CC experiments. It is also used in gene therapy. Osterix protein is useful
 CC for controlling bone formation, serves as receptors of soluble molecules,
 CC e.g. cytokines, growth factors, etc, as homing/adhesion/ rolling
 CC receptors mediating the migration of osteoblast, as signalling receptors,
 CC thereby regulating function of osteoblasts, and/or ligands of signalling
 CC receptors or osteoblast. It is also useful for purifying osteoblasts,
 CC from a test composition suspected of containing the cells. The present
 CC sequence is mouse osterix cDNA. Mouse osterix gene is located at
 CC chromosome 15
 XX
 SQ Sequence 2960 BP; 647 A; 873 C; 719 G; 721 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.51e-87	Length:	2960
Score:	2384.00	Matches:	428
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-09-734-329B-2 (1-428) x AAD40709 (1-2960)

Qy	1	MetAlaSerSerLeuLeuGluGluGluAlaHisTyrGlySerSerProLeuAlaMetLeu	20
Db	100	ATGGCGTCCTCTCTGCTTGAGGAAGAAGCTCACTATGGCTCCAGTCCCCTGGCCATGCTG	159
Qy	21	ThrAlaAlaCysSerLysPheGlyGlySerSerProLeuArgAspSerThrThrLeuGly	40
Db	160	ACTGCAGCCTGCAGCAAATTTGGCGGCTCTAGCCCTCTGCGGGACTCAACAACCCTGGGG	219
Qy	41	LysGlyGlyThrLysLysProTyrAlaAspLeuSerAlaProLysThrMetGlyAspAla	60
Db	220	AAAGGAGGCACAAAGAAGCCATACGCTGACCTTTCAGCCCCCAAACCATGGGGGACGCC	279
Qy	61	TyrProAlaProPheSerSerThrAsnGlyLeuLeuSerProAlaGlySerProProAla	80
Db	280	TACCCAGCTCCCTTCTCAAGCACCAATGGACTCCTCTCTCCTGCAGGCAGTCTCCGGCC	339
Qy	81	ProAlaSerGlyTyrAlaAsnAspTyrProProPheProHisSerPheProGlyProThr	100
Db	340	CCAGCCTCTGGCTATGCAAATGACTACCCACCCTTCCCTCACTCATTTCCTGGGCCCACC	399
Qy	101	GlyAlaGlnAspProGlyLeuLeuValProLysGlyHisSerSerSerAspCysLeuPro	120
Db	400	GGTGGCCCAAGACCCTGGGCTCCTAGTGCCTAAGGGGCACAGCTCGTCTGACTGCCTGCCT	459
Qy	121	SerValTyrThrSerLeuAspMetThrHisProTyrGlySerTrpTyrLysAlaGlyIle	140
Db	460	AGTGTCTACACTTCCCTGGATATGACTCATCCCTATGGCTCGTGGTACAAGGCAGGCATC	519
Qy	141	HisAlaGlyIleSerProGlyProGlyAsnThrProThrProTrpTrpAspMetHisPro	160
Db	520	CACGCAGGCATCTCACCAGGTCCAGGCAACACACCTACTCCTTGGTGGGACATGCACCCT	579
Qy	161	GlyGlyAsnTrpLeuGlyGlyGlyGlnGlyGlnGlyAspGlyLeuGlnGlyThrLeuSer	180
Db	580	GGGGGCAACTGGCTAGGTGGTGGTCAGGGCCAGGGTGATGGGCTGCAAGGGACACTGTCC	639

Qy	181	ThrGlyProAlaGlnProProLeuAsnProGlnLeuProThrTyrProSerAspPheAla	200
Db	640	ACAGGCCCTGCCAGCCTCCACTGAACCCCCAGCTGCCTACTTACCCATCTGACTTTGCT	699
Qy	201	ProLeuAsnProAlaProTyrProAlaProHisLeuLeuGlnProGlyProGlnHisVal	220
Db	700	CCCCTTAACCCAGCTCCCTACCCAGCGCCCCACCTCTTGCAACCAGGGCCCCAGCATGTC	759
Qy	221	LeuProGlnAspValTyrLysProLysAlaValGlyAsnSerGlyGlnLeuGluGlySer	240
Db	760	CTACCCCAAGATGTC TATAAGCCCAAGGCGGTTGGCAATAGTGGGCAACTGGAGGGGAGT	819
Qy	241	GlyAlaAlaLysProProArgGlyAlaGlyThrGlyGlySerGlyGlyTyrAlaGlySer	260
Db	820	GGTGCAGCCAAACCCCTCGGGGTGCTGGCACAGGGGGCAGCGGTGGATATGCGGGCAGT	879
Qy	261	GlyAlaGlyArgSerThrCysAspCysProAsnCysGlnGluLeuGluArgLeuGlyAla	280
Db	880	GGGGCAGGGCGTTCTACCTGCGACTGCCCAACTGTCAGGAGCTAGAGCGGCTCGGGGCA	939
Qy	281	AlaAlaAlaGlyLeuArgLysLysProIleHisSerCysHisIleProGlyCysGlyLys	300
Db	940	GCAGCGGCTGGGCTGAGGAAGAAGCCCATTACAGCTGCCACATCCCTGGGTGCGGCAAG	999
Qy	301	ValTyrGlyLysAlaSerHisLeuLysAlaHisLeuArgTrpHisThrGlyGluArgPro	320
Db	1000	GTGTACGGCAAGGCTTCGCATCTGAAAGCCCACTTGCGCTGGCACACTGGCGAGAGGCCT	1059
Qy	321	PheValCysAsnTrpLeuPheCysGlyLysArgPheThrArgSerAspGluLeuGluArg	340
Db	1060	TTCGTCTGCAACTGGCTTTTCTGCGGCAAGAGGTTCACTCGCTCTGACGAGCTGGAGCGC	1119
Qy	341	HisValArgThrHisThrArgGluLysLysPheThrCysLeuLeuCysSerLysArgPhe	360
Db	1120	CACGTGCGCACTCACACCCGGGAGAAGAAGTTCACTTGCTTGCTCTGTTCCAAGCGCTTT	1179
Qy	361	ThrArgSerAspHisLeuSerLysHisGlnArgThrHisGlyGluProGlyProGlyPro	380
Db	1180	ACCAGAAGCGACCACCTTGAGCAAACATCAGCGCACCCACGGGGAGCCAGGCCCGGGACCG	1239
Qy	381	ProProSerGlyProLysGluLeuGlyGluGlyArgSerValGlyGluGluGluAlaAsn	400
Db	1240	CCCCAAGTGGCCCTAAGGAGCTGGGGGAGGGTCGCAGCGTCGGGGAAGAAGAAGCCAAT	1299
Qy	401	GlnProProArgSerSerThrSerProAlaProProGluLysAlaHisGlyGlySerPro	420
Db	1300	CAGCCGCCCCGATCTTCCACTTCGCCTGCACCCCCAGAAAAAGCCACGGAGGCAGCCCA	1359
Qy	421	GluGlnSerAsnLeuLeuGluIle	428
Db	1360	GAGCAGAGCAACCTGCTAGAGATC	1383

RESULT 2

ADR08250

ID ADR08250 standard; cDNA; 3153 BP.

XX

AC ADR08250;

XX

DT 04-NOV-2004 (first entry)

XX

SCORE Search Results Details for Application 09734329 and Search Result us-09-734-329b-2.p2n.rge.

[Score Home Page](#)
[Retrieve Application List](#)
[SCORE System Overview](#)
[SCORE FAQ](#)
[Comments / Suggestions](#)

This page gives you Search Results detail for the Application 09734329 and Search Result us-09-734-329b-2.p2n.rge.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2006, 09:19:19 ; Search time 10184 Seconds
(without alignments)
4031.249 Million cell updates/sec

Title: US-09-734-329B-2
Perfect score: 2384
Sequence: 1 MASSLLEEEAHYGSSPLAML.....PAPPEKAHGGSPQSNLLEI 428

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/abss/ABSSWEB_spool/US09734329/runat_25072006_091902_3663/app_query.fasta_1
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss06p
-USER=US09734329 @CGN_1_1_8328 @runat_25072006_091902_3663 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database : GenEmbl:*
1: gb_env:*

2: gb_pat:*
 3: gb_ph:*
 4: gb_pl:*
 5: gb_pr:*
 6: gb_ro:*
 7: gb_sts:*
 8: gb_sy:*
 9: gb_un:*
 10: gb_vi:*
 11: gb_ov:*
 12: gb_htg:*
 13: gb_in:*
 14: gb_om:*
 15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	%		Query		DB	ID	Description
	No.	Score	Match	Length			
	1	2384	100.0	2284	6	BC113150	BC113150 Mus muscu
	2	2384	100.0	2960	2	AX555029	AX555029 Sequence
	3	2384	100.0	2960	6	AF184902	AF184902 Mus muscu
c	4	2360	99.0	138860	12	AC055703	AC055703 Mus muscu
c	5	2360	99.0	205599	6	AC137156	AC137156 Mus muscu
c	6	2360	99.0	208289	6	AC163291	AC163291 Mus muscu
	7	2356	98.8	1403	6	AY803733	AY803733 Mus muscu
	8	2338	98.1	1287	6	AY177399	AY177399 Rattus no
c	9	2333	97.9	254686	12	AC097309	AC097309 Rattus no
	10	2264.5	95.0	1357	5	BC101549	BC101549 Homo sapi
	11	2264.5	95.0	2843	5	AY150673	AY150673 Homo sapi
	12	2264.5	95.0	2971	5	AF466179	AF466179 Homo sapi
	13	2264.5	95.0	2995	5	AF477981	AF477981 Homo sapi
	14	2264.5	95.0	3153	2	CQ851287	CQ851287 Sequence
	15	2264.5	95.0	3153	5	AK128520	AK128520 Homo sapi
	16	2240.5	94.0	1575	2	CQ740107	CQ740107 Sequence
c	17	2239.5	93.9	106551	5	AC073611	AC073611 Homo sapi
	18	2239.5	93.9	166697	12	AC021103	AC021103 Homo sapi
	19	2236.5	93.8	1350	5	AY150674	AY150674 Homo sapi
	20	2212.5	92.8	173487	12	AC166700	AC166700 Bos tauru
	21	1301	54.6	1433	11	AY380818	AY380818 Danio rer
	22	866.5	36.3	2381	11	AY769697	AY769697 Danio rer
	23	857.5	36.0	1940	11	BC067673	BC067673 Danio rer
	24	857.5	36.0	1977	11	AY731230	AY731230 Danio rer
	25	854.5	35.8	2380	6	AY591908	AY591908 Mus muscu
	26	843	35.4	1765	11	AY769696	AY769696 Danio rer
	27	842.5	35.3	2459	11	AY591905	AY591905 Danio rer
	28	842.5	35.3	240532	12	AC130082	AC130082 Rattus no
	29	838.5	35.2	2382	11	BC067654	BC067654 Danio rer
c	30	836.5	35.1	3128	2	CQ851014	CQ851014 Sequence
c	31	836.5	35.1	3128	5	AK128215	AK128215 Homo sapi
	32	836.5	35.1	171734	6	AL954713	AL954713 Mouse DNA
	33	836.5	35.1	174437	5	AC018470	AC018470 Homo sapi
	34	836.5	35.1	211696	12	AC099580	AC099580 Mus muscu
	35	833	34.9	2053	11	AY591904	AY591904 Danio rer
	36	829	34.8	155001	11	BX005032	BX005032 Zebrafish
	37	828.5	34.8	1618	5	AY167047	AY167047 Homo sapi

38	823	34.5	1747	11	BC065597	BC065597 Danio rer
39	822.5	34.5	2919	11	AY457141	AY457141 Danio rer
40	818.5	34.3	3471	2	AX714486	AX714486 Sequence
41	818.5	34.3	3471	5	AK056857	AK056857 Homo sapi
42	815.5	34.2	190372	6	AC131717	AC131717 Mus muscu
43	815.5	34.2	191231	6	AC126277	AC126277 Mus muscu
44	811.5	34.0	3581	5	BC038669	BC038669 Homo sapi
45	810.5	34.0	1606	5	AY167048	AY167048 Homo sapi

ALIGNMENTS

RESULT 1

BC113150

LOCUS BC113150 2284 bp mRNA linear ROD 06-FEB-2006

DEFINITION Mus musculus trans-acting transcription factor 7, mRNA (cDNA clone MGC:130562 IMAGE:40054399), complete cds.

ACCESSION BC113150

VERSION BC113150.1 GI:86577831

KEYWORDS MGC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2284)

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

CONSRMT Mammalian Gene Collection Program Team

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932

REFERENCE 2 (bases 1 to 2284)

AUTHORS .

CONSRMT NIH MGC Project

TITLE Direct Submission

JOURNAL Submitted (05-FEB-2006) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Baylor Human Genome Sequencing Center

cDNA Library Preparation: Baylor Human Genome Sequencing Center

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAM Plate: 15 Row: f Column: 19.

FEATURES	Location/Qualifiers
source	1. .2284 /organism="Mus musculus" /mol_type="mRNA" /db_xref="taxon:10090" /clone="MGC:130562 IMAGE:40054399" /tissue_type="PCR rescued clones" /clone_lib="NIH_MGC_285" /note="Vector: pCR-Blunt II-TOPO with reversed insert; Clone identification sequence tag: ATAGAGTG"
gene	1. .2284 /gene="Sp7" /note="synonym: C22" /db_xref="GeneID:170574" /db_xref="MGI:2153568"
CDS	38. .1324 /gene="Sp7" /codon_start=1 /product="Sp7 protein" /protein_id="AAI13151.1" /db_xref="GI:86577832" /db_xref="GeneID:170574" /db_xref="MGI:2153568" /translation="MASSLLEEEAHYGSS PLAMLTAACSKFGGSS PLRDSTTLGKGGT KKPYADLSAPKTMGDYAPAFSSSTNGLLSPAGSPAPASGYANDY PFFPHSFPGP TGA QDPGLLVPKGHSSSDCLPSVYTS LDMTHPYGSWYKAGIHAGISPGPGNTPT PWDMHP GGNWLGGGQGGDGLQGTLSTGPAQPPLNPQLPTYPSDFAPLNPAPYP APHLLQPGPQ HVL PQDVYKPKAVGN SGQLEGSGAAKPPRGAGTGGSGGYAGSG AGRSTCDCPNCQELE RLGAAAAGLRKKPIHSCHIPGCGKVY GKASHLKAHLRWHT GERPFVCNWLF CGKRFT R SDELERHVRTH TREKKFTCLLCSKRFT RSDHLSKH QRT HGE PGP GPPPSGPKELGEGR SVGEEANQPPRSSTSPAPPEKAHGGSP EQSNLLEI"

ORIGIN

Alignment Scores:

Pred. No.:	3.49e-131	Length:	2284
Score:	2384.00	Matches:	428
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-09-734-329B-2 (1-428) x BC113150 (1-2284)

Qy 1 MetAlaSerSerLeuLeuGluGluGluAlaHisTyrGlySerSerProLeuAlaMetLeu 20

Db	38		ATGGCGTCCTCTCTGCTTGAGGAAGAAGCTCACTATGGCTCCAGTCCCCTGGCCATGCTG	97
Qy	21		ThrAlaAlaCysSerLysPheGlyGlySerSerProLeuArgAspSerThrThrLeuGly	40
Db	98		ACTGCAGCCTGCAGCAAATTTGGCGGCTCTAGCCCTCTGCGGGACTCAACAACCTGGGG	157
Qy	41		LysGlyGlyThrLysLysProTyrAlaAspLeuSerAlaProLysThrMetGlyAspAla	60
Db	158		AAAGGAGGCACAAAGAAGCCATACGCTGACCTTTCAGCCCCAAAACCATGGGGGACGCC	217
Qy	61		TyrProAlaProPheSerSerThrAsnGlyLeuLeuSerProAlaGlySerProProAla	80
Db	218		TACCCAGCTCCCTTCTCAAGCACCAATGGACTCCTCTCTCCTGCAGGCAGTCCCTCCGGCC	277
Qy	81		ProAlaSerGlyTyrAlaAsnAspTyrProProPheProHisSerPheProGlyProThr	100
Db	278		CCAGCCTCTGGCTATGCAAATGACTACCCACCCTTCCCTCACTCATTTCTGGGCCCCACC	337
Qy	101		GlyAlaGlnAspProGlyLeuLeuValProLysGlyHisSerSerSerAspCysLeuPro	120
Db	338		GGTGCCCAAGACCCTGGGCTCCTAGTGCCTAAGGGGCACAGCTCGTCTGACTGCCTGCCT	397
Qy	121		SerValTyrThrSerLeuAspMetThrHisProTyrGlySerTrpTyrLysAlaGlyIle	140
Db	398		AGTGTCTACACTTCCTGGATATGACTCATCCCTATGGCTCGTGGTACAAGGCAGGCATC	457
Qy	141		HisAlaGlyIleSerProGlyProGlyAsnThrProThrProTrpTrpAspMetHisPro	160
Db	458		CACGCAGGCATCTCACCAGGTCCAGGCAACACACCTACTCCTTGGTGGGACATGCACCCT	517
Qy	161		GlyGlyAsnTrpLeuGlyGlyGlyGlnGlyGlnGlyAspGlyLeuGlnGlyThrLeuSer	180
Db	518		GGGGGCAACTGGCTAGGTGGTGGTCAGGGCCAGGGTGATGGGCTGCAAGGGACACTGTCC	577
Qy	181		ThrGlyProAlaGlnProProLeuAsnProGlnLeuProThrTyrProSerAspPheAla	200
Db	578		ACAGGCCCTGCCAGCCTCCACTGAACCCCCAGCTGCCTACTTACCCATCTGACTTTGCT	637
Qy	201		ProLeuAsnProAlaProTyrProAlaProHisLeuLeuGlnProGlyProGlnHisVal	220
Db	638		CCCCTTAACCCAGCTCCCTACCCAGCGCCCCACCTCTTGCAACCAGGGCCCCAGCATGTC	697
Qy	221		LeuProGlnAspValTyrLysProLysAlaValGlyAsnSerGlyGlnLeuGluGlySer	240
Db	698		CTACCCCAAGATGTCTATAAGCCCAAGGCGGTGGCAATAGTGGGCAACTGGAGGGGAGT	757
Qy	241		GlyAlaAlaLysProProArgGlyAlaGlyThrGlyGlySerGlyGlyTyrAlaGlySer	260
Db	758		GGTGCAGCCAAACCCCTCGGGGTGCTGGCACAGGGGGCAGCGGTGATATGCGGGCAGT	817
Qy	261		GlyAlaGlyArgSerThrCysAspCysProAsnCysGlnGluLeuGluArgLeuGlyAla	280
Db	818		GGGGCAGGGCGTTCTACCTGCGACTGCCCAACTGTCAGGAGCTAGAGCGGCTCGGGGCA	877
Qy	281		AlaAlaAlaGlyLeuArgLysLysProIleHisSerCysHisIleProGlyCysGlyLys	300
Db	878		GCAGCGGCTGGGCTGAGGAAGAAGCCCATTCACAGCTGCCACATCCCTGGGTGCGGCAAG	937
Qy	301		ValTyrGlyLysAlaSerHisLeuLysAlaHisLeuArgTrpHisThrGlyGluArgPro	320

```

Db      938 GTGTACGGCAAGGCTTCGCATCTGAAAGCCCACTTGCGCTGGCACACTGGCGAGAGGCCT 997
Qy      321 PheValCysAsnTrpLeuPheCysGlyLysArgPheThrArgSerAspGluLeuGluArg 340
      |||
Db      998 TTCGTCTGCAACTGGCTTTTCTGCGGCAAGAGGTTCACTCGCTCTGACGAGCTGGAGCGC 1057
Qy      341 HisValArgThrHisThrArgGluLysLysPheThrCysLeuLeuCysSerLysArgPhe 360
      |||
Db      1058 CACGTGCGCACTCACACCCGGGAGAAGAAGTTCACCTGCCCTGCTCTGTTCCAAGCGCTTT 1117
Qy      361 ThrArgSerAspHisLeuSerLysHisGlnArgThrHisGlyGluProGlyProGlyPro 380
      |||
Db      1118 ACCAGAAGCGACCACTTGAGCAAACATCAGCGCACCCACGGGGAGCCAGGCCCGGGACCG 1177
Qy      381 ProProSerGlyProLysGluLeuGlyGluGlyArgSerValGlyGluGluGluAlaAsn 400
      |||
Db      1178 CCCCCAAGTGCCCTAAGGAGCTGGGGGAGGGTCGCAGCGTCGGGGAAGAAGAAGCCAAT 1237
Qy      401 GlnProProArgSerSerThrSerProAlaProProGluLysAlaHisGlyGlySerPro 420
      |||
Db      1238 CAGCCGCCCGATCTTCCACTTCGCCTGCACCCCAGAAAAAGCCACGGAGGCAGCCCA 1297
Qy      421 GluGlnSerAsnLeuLeuGluIle 428
      |||
Db      1298 GAGCAGAGCAACCTGCTAGAGATC 1321

```

RESULT 2

AX555029

LOCUS AX555029 2960 bp DNA linear PAT 27-NOV-2002

DEFINITION Sequence 1 from Patent WO0244380.

ACCESSION AX555029

VERSION AX555029.1 GI:25898589

KEYWORDS .

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS de Crombrughe,B., Nakashima,K. and Zhou,X.

TITLE Master bone formation transcription factor: compositions and methods of use

JOURNAL Patent: WO 0244380-A 1 06-JUN-2002;

BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)

FEATURES Location/Qualifiers

source 1. .2960

/organism="Mus musculus"

/mol_type="unassigned DNA"

/db_xref="taxon:10090"

CDS 100. .1386

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD58224.1"

/db_xref="GI:25898590"

/translation="MASSLLEEEAHYGSSPLAMLTAACSKFGGSSPLRDSTTLGKGGT
KKPYADLSAPKTMGDAYPAPFSSTNGLLSPAGSPAPASGYANDYPFPHSFPGPPTGA
QDPGLLVPKGHSSSDCLPSVYTSLDMTHPYGSWYKAGIHAGISPGPGNTPTPWDMHP
GGNWLGGGQGGDGLQGTLSGPAQPPLNPQLPTYPSDFAPLNPAPYPAPHLLQPGPQ
HVLPPQDVYKPKAVGNSGQLEGSGAAKPPRGAGTGGSGGYAGSGAGRSTCDCPNCQELE
RLGAAAAGLRKKPIHSCHIPGCGKVYGKASHLKAHLRWHTGERPFVCNWLFCKGRFTR
SDELERHVRTHTREKKFTCLLCSKRFTSRDHLKHKQRTHEPGPGPPSPGPKELGEGR

SVGEEENQPPRSSTSPAPPEKAHGGSP EQSNLLEI"

ORIGIN

Alignment Scores:

Pred. No.:	4.78e-131	Length:	2960
Score:	2384.00	Matches:	428
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

US-09-734-329B-2 (1-428) x AX555029 (1-2960)

Qy	1	MetAlaSerSerLeuLeuGluGluGluAlaHisTyrGlySerSerProLeuAlaMetLeu	20
Db	100	ATGGCGTCCTCTCTGCTTGAGGAAGAAGCTCACTATGGCTCCAGTCCCTTGCCCATGCTG	159
Qy	21	ThrAlaAlaCysSerLysPheGlyGlySerSerProLeuArgAspSerThrThrLeuGly	40
Db	160	ACTGCAGCCTGCAGCAAATTTGGCGGCTCTAGCCCTCTGCGGGACTCAACAACCTGGGG	219
Qy	41	LysGlyGlyThrLysLysProTyrAlaAspLeuSerAlaProLysThrMetGlyAspAla	60
Db	220	AAAGGAGGCACAAAGAAGCCATACGCTGACCTTTCAGCCCCAAAACCATGGGGGACGCC	279
Qy	61	TyrProAlaProPheSerSerThrAsnGlyLeuLeuSerProAlaGlySerProProAla	80
Db	280	TACCCAGCTCCCTTCTCAAGCACCAATGGACTCCTCTCTCCTGCAGGCAGTCTCCGGCC	339
Qy	81	ProAlaSerGlyTyrAlaAsnAspTyrProProPheProHisSerPheProGlyProThr	100
Db	340	CCAGCCTCTGGCTATGCAAATGACTACCCACCCTTCCTCACTCATTTCTGGGCCCACC	399
Qy	101	GlyAlaGlnAspProGlyLeuLeuValProLysGlyHisSerSerSerAspCysLeuPro	120
Db	400	GGTGCCCAAGACCCTGGGCTCCTAGTGCCTAAGGGGCACAGCTCGTCTGACTGCCTGCCT	459
Qy	121	SerValTyrThrSerLeuAspMetThrHisProTyrGlySerTrpTyrLysAlaGlyIle	140
Db	460	AGTGTCTACACTTCCCTGGATATGACTCATCCCTATGGCTCGTGGTACAAGGCAGGCATC	519
Qy	141	HisAlaGlyIleSerProGlyProGlyAsnThrProThrProTrpTrpAspMetHisPro	160
Db	520	CACGCAGGCATCTACCAGGTCCAGGCAACACACCTACTCCTTGGTGGGACATGCACCCT	579
Qy	161	GlyGlyAsnTrpLeuGlyGlyGlyGlnGlyGlnGlyAspGlyLeuGlnGlyThrLeuSer	180
Db	580	GGGGGCAACTGGCTAGGTGGTGGTCAGGGCCAGGGTGATGGGCTGCAAGGGACACTGTCC	639
Qy	181	ThrGlyProAlaGlnProProLeuAsnProGlnLeuProThrTyrProSerAspPheAla	200
Db	640	ACAGGCCCTGCCCAGCCTCCACTGAACCCCAGCTGCCTACTTACCCATCTGACTTTGCT	699
Qy	201	ProLeuAsnProAlaProTyrProAlaProHisLeuLeuGlnProGlyProGlnHisVal	220
Db	700	CCCCTTAACCCAGCTCCCTACCCAGCGCCACCTCTTGCAACCAGGGCCCCAGCATGTC	759
Qy	221	LeuProGlnAspValTyrLysProLysAlaValGlyAsnSerGlyGlnLeuGluGlySer	240
Db	760	CTACCCCAAGATGTCTATAAGCCCAAGGCGGTTGGCAATAGTGGGCAACTGGAGGGGAGT	819

```

Qy      241 GlyAlaAlaLysProProArgGlyAlaGlyThrGlyGlySerGlyGlyTyrAlaGlySer 260
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      820 GGTGCAGCCAAACCCCTCGGGGTGCTGGCACAGGGGGCAGCGGTGGATATGCGGGCAGT 879

Qy      261 GlyAlaGlyArgSerThrCysAspCysProAsnCysGlnGluLeuGluArgLeuGlyAla 280
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      880 GGGGCAGGGCGTTCTACCTGCGACTGCCCCAACTGTCAGGAGCTAGAGCGGCTCGGGGCA 939

Qy      281 AlaAlaAlaGlyLeuArgLysLysProIleHisSerCysHisIleProGlyCysGlyLys 300
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      940 GCAGCGGCTGGGCTGAGGAAGAAGCCCATTACAGCTGCCACATCCCTGGGTGCGGCAAG 999

Qy      301 ValTyrGlyLysAlaSerHisLeuLysAlaHisLeuArgTrpHisThrGlyGluArgPro 320
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1000 GTGTACGGCAAGGCTTCGCATCTGAAAGCCCACTTGCGCTGGCACACTGGCGAGAGGCCT 1059

Qy      321 PheValCysAsnTrpLeuPheCysGlyLysArgPheThrArgSerAspGluLeuGluArg 340
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1060 TTCGTCTGCAACTGGCTTTTCTGCGGCAAGAGGTTCACTCGCTCTGACGAGCTGGAGCGC 1119

Qy      341 HisValArgThrHisThrArgGluLysLysPheThrCysLeuLeuCysSerLysArgPhe 360
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1120 CACGTGCGCACTCACACCCGGGAGAAGAAGTTCCTTGCCTGCTCTGTTCCAAGCGCTTT 1179

Qy      361 ThrArgSerAspHisLeuSerLysHisGlnArgThrHisGlyGluProGlyProGlyPro 380
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1180 ACCAGAAGCGACCACCTTGAGCAAACATCAGCGCACCCACGGGGAGCCAGGCCCCGGGACCG 1239

Qy      381 ProProSerGlyProLysGluLeuGlyGluGlyArgSerValGlyGluGluGluAlaAsn 400
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1240 CCCCCAAGTGGCCCTAAGGAGCTGGGGGAGGGTCGCAGCGTCGGGGAAGAAGAAGCCAAT 1299

Qy      401 GlnProProArgSerSerThrSerProAlaProProGluLysAlaHisGlyGlySerPro 420
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1300 CAGCCGCCCCGATCTTCCACTTCGCCTGCACCCCCAGAAAAGCCACGGAGGCAGCCCA 1359

Qy      421 GluGlnSerAsnLeuLeuGluIle 428
      ||||||||||||||||||||
Db     1360 GAGCAGAGCAACCTGCTAGAGATC 1383

```

RESULT 3

```

AF184902
LOCUS      AF184902                2960 bp    mRNA    linear    ROD 12-JAN-2002
DEFINITION Mus musculus Osterix (C22) mRNA, complete cds.
ACCESSION  AF184902
VERSION    AF184902.1  GI:18138080
KEYWORDS   .
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 2960)
AUTHORS    Nakashima,K., Zhou,X., Kunkel,G., Zhang,Z., Deng,J.M.,
            Behringer,R.R. and de Crombrughe,B.
TITLE      The novel zinc finger-containing transcription factor osterix is
            required for osteoblast differentiation and bone formation
JOURNAL    Cell 108 (1), 17-29 (2002)
PUBMED     11792318
REFERENCE  2 (bases 1 to 2960)
AUTHORS    Nakashima,K., Zhou,X. and de Crombrughe,B.

```

TITLE Direct Submission
JOURNAL Submitted (10-SEP-1999) Molecular Genetics, M.D. Anderson Cancer Center, 1515 Holcombe, Houston, TX 77030, USA

FEATURES Location/Qualifiers
source 1. .2960
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/cell_line="C2C12"
/cell_type="myoblasts"
gene 1. .2960
/gene="C22"
CDS 100. .1386
/gene="C22"
/note="zinc finger protein; osteoblast/chondrocyte specific transcription factor"
/codon_start=1
/product="Osterix"
/protein_id="AAL60067.1"
/db_xref="GI:18138081"
/translation="MASSLLEEEAHYGSSPLAMLTAAACSKFGGSSPLRDSTTLGKGGT
KKPYADLSAPKTMGDYAPFSSSTNGLLS PAPS PPAPASGYANDYPPFPHSFPGPTGA
QDPGLLVPKGHSSSDCLPSVYTSLDMTHPYGSWYKAGIHAGISPGPGNTPTPWDMHP
GGNWLGQGGQGDGLQGTLSGTGPAQPPLNPQLPTYPSDFAPLNPAPYPAPHLLOPGPQ
HVLPPQDVYKPKAVGNSGQLEGSGAAKPPRGAGTGGSGGYAGSGAGRSTCDCPNCQELE
RLGAAAAGLRKKPIHSCHIPGCGKVYKGASHLKAHLRWHTGERPFVCNWLFCEGKRFRTR
SDELERHVRTHTREKKFTCLLC SKRFRTRSDHLSKHQRTHGEPGPGPPPSGPKELGEGR
SVGEEEAQPPRSSTSPAPPEKAHGGSP EQSNLLEI"

ORIGIN

Alignment Scores:

Pred. No.:	4.78e-131	Length:	2960
Score:	2384.00	Matches:	428
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-09-734-329B-2 (1-428) x AF184902 (1-2960)

Qy	1	MetAlaSerSerLeuLeuGluGluGluAlaHisTyrGlySerSerProLeuAlaMetLeu	20
Db	100	ATGGCGTCTCTGCTTGAGGAAGAAGCTCACTATGGCTCCAGTCCCCTGGCCATGCTG	159
Qy	21	ThrAlaAlaCysSerLysPheGlyGlySerSerProLeuArgAspSerThrThrLeuGly	40
Db	160	ACTGCAGCCTGCAGCAAATTTGGCGGCTCTAGCCCTCTGCGGGACTCAACAACCCTGGGG	219
Qy	41	LysGlyGlyThrLysLysProTyrAlaAspLeuSerAlaProLysThrMetGlyAspAla	60
Db	220	AAAGGAGGCACAAAGAAGCCATACGCTGACCTTTCAGCCCCAAACCATGGGGGACGCC	279
Qy	61	TyrProAlaProPheSerSerThrAsnGlyLeuLeuSerProAlaGlySerProProAla	80
Db	280	TACCCAGCTCCCTTCTCAAGCACCAATGGACTCCTCTCTCCTGCAGGCAGTCTCCGGCC	339
Qy	81	ProAlaSerGlyTyrAlaAsnAspTyrProProPheProHisSerPheProGlyProThr	100
Db	340	CCAGCCTCTGGCTATGCAAATGACTACCCACCCTTCCCTCACTCATTTCTGGGGCCACC	399
Qy	101	GlyAlaGlnAspProGlyLeuLeuValProLysGlyHisSerSerSerAspCysLeuPro	120

Db	400		GGTGCCCAAGACCCTGGGCTCCTAGTGCCTAAGGGGCACAGCTCGTCTGACTGCCTGCCT	459
Qy	121		SerValTyrThrSerLeuAspMetThrHisProTyrGlySerTrpTyrLysAlaGlyIle	140
Db	460		AGTGTCTACACTTCCTGGATATGACTCATCCCTATGGCTCGTGGTACAAGGCAGGCATC	519
Qy	141		HisAlaGlyIleSerProGlyProGlyAsnThrProThrProTrpTrpAspMetHisPro	160
Db	520		CACGCAGGCATCTCACCAGGTCCAGGCAACACACCTACTCCTTGGTGGGACATGCACCT	579
Qy	161		GlyGlyAsnTrpLeuGlyGlyGlyGlnGlyGlnGlyAspGlyLeuGlnGlyThrLeuSer	180
Db	580		GGGGGCAACTGGCTAGGTGGTGGTCAGGGCCAGGGTGATGGGCTGCAAGGGACACTGTCC	639
Qy	181		ThrGlyProAlaGlnProProLeuAsnProGlnLeuProThrTyrProSerAspPheAla	200
Db	640		ACAGGCCCTGCCAGCCTCCACTGAACCCCCAGCTGCCTACTTACCCATCTGACTTTGCT	699
Qy	201		ProLeuAsnProAlaProTyrProAlaProHisLeuLeuGlnProGlyProGlnHisVal	220
Db	700		CCCCTTAACCCAGCTCCCTACCCAGCGCCCCACCTCTTGCAACCAGGGCCCCAGCATGTC	759
Qy	221		LeuProGlnAspValTyrLysProLysAlaValGlyAsnSerGlyGlnLeuGluGlySer	240
Db	760		CTACCCCAAGATGTCTATAAGCCCAAGGCGGTTGGCAATAGTGGGCAACTGGAGGGGAGT	819
Qy	241		GlyAlaAlaLysProProArgGlyAlaGlyThrGlyGlySerGlyGlyTyrAlaGlySer	260
Db	820		GGTGCAGCCAAACCCCTCGGGGTGCTGGCACAGGGGGCAGCGGTGGATATGCGGGCAGT	879
Qy	261		GlyAlaGlyArgSerThrCysAspCysProAsnCysGlnGluLeuGluArgLeuGlyAla	280
Db	880		GGGGCAGGGCGTTCTACCTGCGACTGCCCAACTGTCAGGAGCTAGAGCGGCTCGGGGCA	939
Qy	281		AlaAlaAlaGlyLeuArgLysLysProIleHisSerCysHisIleProGlyCysGlyLys	300
Db	940		GCAGCGGCTGGGCTGAGGAAGAAGCCATTACAGCTGCCACATCCCTGGGTGCGGCAAG	999
Qy	301		ValTyrGlyLysAlaSerHisLeuLysAlaHisLeuArgTrpHisThrGlyGluArgPro	320
Db	1000		GTGTACGGCAAGGCTTCGCATCTGAAAGCCACTTGCCTGCGCTGGCACACTGGCGAGAGGCCT	1059
Qy	321		PheValCysAsnTrpLeuPheCysGlyLysArgPheThrArgSerAspGluLeuGluArg	340
Db	1060		TTCGTCTGCAACTGGCTTTTCTGCGGCAAGAGGTTCACTCGCTCTGACGAGCTGGAGCGC	1119
Qy	341		HisValArgThrHisThrArgGluLysLysPheThrCysLeuLeuCysSerLysArgPhe	360
Db	1120		CACGTGCGCACTCACACCCGGGAGAAGAAGTTCACTTGCTGCTCTGTTCCAAGCGCTTT	1179
Qy	361		ThrArgSerAspHisLeuSerLysHisGlnArgThrHisGlyGluProGlyProGlyPro	380
Db	1180		ACCAGAAGCGACCACTTGAGCAAACATCAGCGCACCCACGGGGAGCCAGGCCCGGGACCG	1239
Qy	381		ProProSerGlyProLysGluLeuGlyGluGlyArgSerValGlyGluGluGluAlaAsn	400
Db	1240		CCCCAAGTGGCCCTAAGGAGCTGGGGGAGGGTCGCAGCGTCGGGAAGAAGAAGCCAAT	1299
Qy	401		GlnProProArgSerSerThrSerProAlaProProGluLysAlaHisGlyGlySerPro	420

Db 1300 CAGCCGCCCCGATCTTCCACTTCGCCTGCACCCCCAGAAAAAGCCCACGGAGGCAGCCCA 1359
 Qy 421 GluGlnSerAsnLeuLeuGluIle 428
 |||||
 Db 1360 GAGCAGAGCAACCTGCTAGAGATC 1383

RESULT 4

AC055703/c

LOCUS AC055703 138860 bp DNA linear HTG 15-MAY-2002

DEFINITION Mus musculus strain C57BL6/J clone RP23-399N14, WORKING DRAFT
 SEQUENCE, 37 unordered pieces.

ACCESSION AC055703

VERSION AC055703.9 GI:20279384

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 138860)

AUTHORS Grills,G., Han,J., Montgomery,K.T., Lee,E., Long,J., Pomerantz,R.,
 Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,
 Gordon,M., Goltz,J.S. and Kucherlapati,R.

TITLE High Throughput Mouse Sequencing

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 138860)

AUTHORS Grills,G., Han,J., Montgomery,K.T., Lee,E., Long,J., Pomerantz,R.,
 Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,
 Gordon,M., Goltz,J.S. and Kucherlapati,R.

TITLE Direct Submission

JOURNAL Submitted (18-APR-2000) Department of Molecular Genetics, Albert
 Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
 Bronx, NY 10461, USA

REFERENCE 3 (bases 1 to 138860)

AUTHORS Grills,G., Han,J., Montgomery,K.T., Lee,E., Long,J., Pomerantz,R.,
 Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,
 Gordon,M., Goltz,J.S. and Kucherlapati,R.

TITLE Direct Submission

JOURNAL Submitted (24-APR-2002) Harvard Partners Center for Genetics and
 Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA
 02139, USA

COMMENT On Apr 24, 2002 this sequence version replaced gi:18151001.

-----Genome Center

Center: Harvard Partners Genome Center

Center Code: HPGC

Web site: <http://www.hpcgg.org/Sequence/mouse.html>

Contact: hpgc@mendel.mgh.harvard.edu

-----Summary Statistics

Center project name: AAS

Sequencing vector: pUC18; L08752

Chemistry: Dye-terminator Big Dye; 100%

*Consensus quality: 126666 at least Q20

*Consensus quality: 122952 at least Q30

*Consensus quality: 116499 at least Q40

Estimated insert size: agarose-FP - N/A

**Estimated insert size: 138140 - sum-of-contigs

Quality coverage: agarose-FP - N/A

Quality coverage: 9.1 x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently

* consists of 37 contigs. The true order of the pieces

SCORE Search Results Details for Application 09734329 and Search Result us-09-734-329b-2.p2n.rst.

[Score Home Page](#)
[Retrieve Application List](#)
[SCORE System Overview](#)
[SCORE FAQ](#)
[Comments / Suggestions](#)

This page gives you Search Results detail for the Application 09734329 and Search Result us-09-734-329b-2.p2n.rst.

[start](#)

[Go Back to previous page](#)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 25, 2006, 09:19:18 ; Search time 6389 Seconds
(without alignments)
5619.066 Million cell updates/sec

Title: US-09-734-329B-2
Perfect score: 2384
Sequence: 1 MASSLLEEEAHYGSSPLAML.....PAPPEKAHGGSPSEQSNLLEI 428

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/abss/ABSSWEB_spool/US09734329/runat_25072006_091905_3703/app_query.fasta_1
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss08
-USER=US09734329 @CGN_1_1_6323 @runat_25072006_091905_3703 -NCPU=6 -ICPU=3
-NO_MMAP -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*

2: gb_est3:*
 3: gb_est4:*
 4: gb_est5:*
 5: gb_est6:*
 6: gb_htc:*
 7: gb_est2:*
 8: gb_est7:*
 9: gb_est8:*
 10: gb_est9:*
 11: gb_gss1:*
 12: gb_gss2:*
 13: gb_gss3:*
 14: gb_gss4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	2384	100.0	2907	6	AK032521	AK032521 Mus muscu
2	2350	98.6	2874	6	AK077375	AK077375 Mus muscu
3	2326	97.6	2917	6	AK076229	AK076229 Mus muscu
4	1182.5	49.6	938	4	BX369746	BX369746 BX369746
5	1119.5	47.0	1014	4	BX394280	BX394280 BX394280
6	1115	46.8	632	5	CF905412	CF905412 A0434B03-
7	1115	46.8	656	5	CF905401	CF905401 A0434A03-
8	1100	46.1	608	5	CF906401	CF906401 A0446H03-
9	1037.5	43.5	745	8	CN287220	CN287220 170005325
10	1000	41.9	615	7	BB622399	BB622399 BB622399
11	808.5	33.9	4301	6	AK030745	AK030745 Mus muscu
12	807.5	33.9	3659	6	AK053093	AK053093 Mus muscu
13	799	33.5	2590	6	AK135123	AK135123 Mus muscu
c 14	754.5	31.6	1798	6	BC023224	BC023224 Mus muscu
15	747.5	31.4	643	7	BB622895	BB622895 BB622895
16	704	29.5	670	7	BB618195	BB618195 BB618195
17	690	28.9	1141	10	DW031048	DW031048 CFW209-B0
c 18	672.5	28.2	5272	6	AK029830	AK029830 Mus muscu
19	667.5	28.0	1131	14	DQ047750	DQ047750 Homo sapi
20	662.5	27.8	1131	14	DQ047751	DQ047751 Pan trogl
21	662	27.8	3599	6	AK137686	AK137686 Mus muscu
22	660	27.7	583	8	CO778267	CO778267 BL003B_F0
23	658.5	27.6	763	3	BP679532	BP679532 BP679532
24	649	27.2	871	8	CX311164	CX311164 JGI_XZT11
25	612.5	25.7	819	5	CK143124	CK143124 AGENCOURT
c 26	608.5	25.5	1088	10	DW665598	DW665598 CNB332-C0
27	606.5	25.4	989	3	BU528562	BU528562 AGENCOURT
28	588.5	24.7	561	3	BQ419249	BQ419249 faa36e03.
c 29	587	24.6	638	1	AL662349	AL662349 AL662349
30	586	24.6	801	9	CX836896	CX836896 JGI_CAAK7
31	576	24.2	2078	3	BP525450	BP525450 BP525450
c 32	572	24.0	641	10	DT247674	DT247674 JGI_CAAU2
33	564.5	23.7	778	12	CC862573	CC862573 NDL.22K1.
34	562	23.6	753	10	DW556033	DW556033 EST_ssal_
35	557.5	23.4	443	1	AL960681	AL960681 AL960681
36	555.5	23.3	474	2	BI979366	BI979366 ft86e04.y
37	553.5	23.2	2097	6	AK036234	AK036234 Mus muscu
38	552	23.2	471	12	CE281682	CE281682 tigr-gss-

c	39	517	21.7	686	10	DW253992	DW253992 UI-S-GB1-
	40	510	21.4	664	2	BJ626154	BJ626154 BJ626154
	41	509	21.4	1084	10	DW665599	DW665599 CNB332-C0
	42	506	21.2	899	5	CD757648	CD757648 AGENCOURT
c	43	504	21.1	694	10	DW271862	DW271862 UI-S-GS1-
	44	504	21.1	866	5	CD755971	CD755971 AGENCOURT
	45	503	21.1	836	5	CD756314	CD756314 AGENCOURT

ALIGNMENTS

RESULT 1

AK032521

LOCUS AK032521 2907 bp mRNA linear HTC 02-SEP-2005

DEFINITION Mus musculus adult male olfactory brain cDNA, RIKEN full-length enriched library, clone:6430578P22 product:osterix, full insert sequence.

ACCESSION AK032521

VERSION AK032521.1 GI:26328334

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

PUBMED 10349636

REFERENCE 2

AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

PUBMED 11076861

REFERENCE 4

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE 5

AUTHORS The FANTOM Consortium, the RIKEN Genome Exploration Research Group Phase I and II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6

AUTHORS RIKEN Genome Exploration Research Group, Genome Science Group
(Genome Network Core Team) and the FANTOM Consortium.

TITLE Antisense Transcription in the Mammalian Transcriptome

JOURNAL Science 309, 1564-1566 (2005)

REFERENCE 7

AUTHORS The FANTOM Consortium, Riken Genome Exploration Research Group and
Genome Science Group (Genome Network Project Core Group).

TITLE The Transcriptional Landscape of the Mammalian Genome

JOURNAL Science 309, 1559-1563 (2005)

REFERENCE 8 (bases 1 to 2907)

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa, 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.

FEATURES Location/Qualifiers

source 1. .2907
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:6430578P22"
/db_xref="taxon:10090"
/clone="6430578P22"
/sex="male"
/tissue_type="olfactory brain"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"

CDS 93. .1379
/note="unnamed protein product; osterix (LocusLink|170574
GB|NM_130458, evidence: BLASTN, 99%, match=1759)
putative"
/codon_start=1
/protein_id="BAC27908.1"
/db_xref="GI:26328335"
/translation="MASSLLEEEAHYGSSPLAMLTAAACKFGGSSPLRDSTTLGKGGT
KKPYADLSAPKTMGDYAPAFSSSTNGLLSPAGSPAPASGYANDYPPFPHSFPGPTGA
QDPGLLVPKGHSSSDCLPSVYTSMDMTHPYGWSYKAGIHAGISPGPGNTPTPWDMHP
GGNWLGGGQGGDGLQGTLSTGPAQPPLNPQLPTYPSTDFAPLNPAYPAPHLLQPGPQ

```

HVLPPQDVYKPKAVGNSGQLEGSGAAKPPRGAGTGGSGGYAGSGAGRSTCDCPNCQELE
RLGAAAAGLRKKPIHSCHIPGCGKVYGKASHLKAHLRWHTGERPFVCNWLFCGKRFTTR
SDELERHVRTHTREKKFTCLLCSKRFTSRDHLSKHQRTTHGEPGPGPPPSGPKELGEGR
SVGEEEAQPPRSSTSPAPPEKAHGGSPSQSNLLEI"
polyA_signal 2891. .2896
               /note="putative"
polyA_site    2907
               /note="putative"

```

ORIGIN

Alignment Scores:

Pred. No.:	3e-143	Length:	2907
Score:	2384.00	Matches:	428
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	6	Gaps:	0

US-09-734-329B-2 (1-428) x AK032521 (1-2907)

Qy	1	MetAlaSerSerLeuLeuGluGluGluAlaHisTyrGlySerSerProLeuAlaMetLeu	20
Db	93	ATGGCGTCCTCTCTGCTTGAGGAAGAAGCTCACTATGGCTCCAGTCCCCTGGCCATGCTG	152
Qy	21	ThrAlaAlaCysSerLysPheGlyGlySerSerProLeuArgAspSerThrThrLeuGly	40
Db	153	ACTGCAGCCTGCAGCAAATTTGGCGGCTCTAGCCCTCTGCGGGACTCAACAACCCTGGGG	212
Qy	41	LysGlyGlyThrLysLysProTyrAlaAspLeuSerAlaProLysThrMetGlyAspAla	60
Db	213	AAAGGAGGCACAAAGAAGCCATACGCTGACCTTTCAGCCCCCAAACCATGGGGGACGCC	272
Qy	61	TyrProAlaProPheSerSerThrAsnGlyLeuLeuSerProAlaGlySerProProAla	80
Db	273	TACCCAGCTCCCTTCTCAAGCACCAATGGACTCCTCTCTCCTGCAGGCAGTCCTCCGGCC	332
Qy	81	ProAlaSerGlyTyrAlaAsnAspTyrProProPheProHisSerPheProGlyProThr	100
Db	333	CCAGCCTCTGGCTATGCAAATGACTACCCACCCTTCCTCACTCATTTCTGGGCCCACC	392
Qy	101	GlyAlaGlnAspProGlyLeuLeuValProLysGlyHisSerSerSerAspCysLeuPro	120
Db	393	GGTGGCCCAAGACCCTGGGCTCCTAGTGCCTAAGGGGCACAGCTCGTCTGACTGCCTGCCT	452
Qy	121	SerValTyrThrSerLeuAspMetThrHisProTyrGlySerTrpTyrLysAlaGlyIle	140
Db	453	AGTGTCTACACTTCCCTGGATATGACTCATCCCTATGGCTCGTGGTACAAGGCAGGCATC	512
Qy	141	HisAlaGlyIleSerProGlyProGlyAsnThrProThrProTrpTrpAspMetHisPro	160
Db	513	CACGCAGGCATCTCACCAGGTCCAGGCAACACACCTACTCCTTGGTGGGACATGCACCCT	572
Qy	161	GlyGlyAsnTrpLeuGlyGlyGlyGlnGlyGlnGlyAspGlyLeuGlnGlyThrLeuSer	180
Db	573	GGGGGCAACTGGCTAGGTGGTGGTCAGGGCCAGGGTGATGGGCTGCAAGGGACACTGTCC	632
Qy	181	ThrGlyProAlaGlnProProLeuAsnProGlnLeuProThrTyrProSerAspPheAla	200
Db	633	ACAGGCCCTGCCCAGCCTCCACTGAACCCCAGCTGCCTACTTACCCATCTGACTTTGCT	692
Qy	201	ProLeuAsnProAlaProTyrProAlaProHisLeuLeuGlnProGlyProGlnHisVal	220

```

      |||
Db      693 CCCCTTAACCCAGCTCCCTACCCAGCGCCCCACCTCTTGCAACCAGGGCCCCAGCATGTC 752
Qy      221 LeuProGlnAspValTyrLysProLysAlaValGlyAsnSerGlyGlnLeuGluGlySer 240
      |||
Db      753 CTACCCCAAGATGTCATAAGCCCAAGGCGGTGGCAATAGTGGGCAACTGGAGGGGAGT 812
Qy      241 GlyAlaAlaLysProProArgGlyAlaGlyThrGlyGlySerGlyGlyTyrAlaGlySer 260
      |||
Db      813 GGTGCAGCCAAACCCCTCGGGGTGCTGGCACAGGGGGCAGCGGTGGATATGCGGGCAGT 872
Qy      261 GlyAlaGlyArgSerThrCysAspCysProAsnCysGlnGluLeuGluArgLeuGlyAla 280
      |||
Db      873 GGGGCAGGGCGTTCTACCTGCGACTGCCCCAACTGTCAGGAGCTAGAGCGGCTCGGGGCA 932
Qy      281 AlaAlaAlaGlyLeuArgLysLysProIleHisSerCysHisIleProGlyCysGlyLys 300
      |||
Db      933 GCAGCGGCTGGGCTGAGGAAGAAGCCATTACAGCTGCCACATCCCTGGGTGCGGCAAG 992
Qy      301 ValTyrGlyLysAlaSerHisLeuLysAlaHisLeuArgTrpHisThrGlyGluArgPro 320
      |||
Db      993 GTGTACGGCAAGGCTTCGCATCTGAAAGCCCACTTGCGCTGGCACACTGGCGAGAGGCCT 1052
Qy      321 PheValCysAsnTrpLeuPheCysGlyLysArgPheThrArgSerAspGluLeuGluArg 340
      |||
Db      1053 TTCGTCTGCAACTGGCTTTTCTGCGGCAAGAGGTTCACTCGCTCTGACGAGCTGGAGCGC 1112
Qy      341 HisValArgThrHisThrArgGluLysLysPheThrCysLeuLeuCysSerLysArgPhe 360
      |||
Db      1113 CACGTGCGCACTCACACCCGGGAGAAGAAGTTCACCTGCCTGCTCTGTTCCAAGCGCTTT 1172
Qy      361 ThrArgSerAspHisLeuSerLysHisGlnArgThrHisGlyGluProGlyProGlyPro 380
      |||
Db      1173 ACCAGAAGCGACCACTTGAGCAAACATCAGCGCACCCACGGGGAGCCAGGCCCGGGACCG 1232
Qy      381 ProProSerGlyProLysGluLeuGlyGluGlyArgSerValGlyGluGluGluAlaAsn 400
      |||
Db      1233 CCCCCAAGTGGCCCTAAGGAGCTGGGGGAGGGTCGCAGCGTCGGGGAAGAAGAAGCCAAT 1292
Qy      401 GlnProProArgSerSerThrSerProAlaProProGluLysAlaHisGlyGlySerPro 420
      |||
Db      1293 CAGCCGCCCCGATCTTCCACTTCGCCTGCACCCCCAGAAAAAGCCCACGGAGGCAGCCCA 1352
Qy      421 GluGlnSerAsnLeuLeuGluIle 428
      |||
Db      1353 GAGCAGAGCAACCTGCTAGAGATC 1376

```

RESULT 2

AK077375

LOCUS AK077375 2874 bp mRNA linear HTC 02-SEP-2005

DEFINITION Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430403B13 product:osterix, full insert sequence.

ACCESSION AK077375

VERSION AK077375.1 GI:26346245

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

REFERENCE 1